斯氏按蚊肽聚糖识别蛋白基因 *PGRP-LC1* 的克隆及功能分析

陈 杨[1,2], 凌尔军[1,*]

(1. 中国科学院上海生命科学研究院植物生理生态研究所昆虫科学研究中心, 上海 200032;
2. 中国科学院研究生院, 北京 100080)

摘要: 天生免疫系统是昆虫抵御外界病原入侵的主要方式。目前研究发现, Imd 信号通路与按蚊感染柏氏疟原虫 *Plasmodium berghei* 的强度密切相关, 而 PGRP-LC1 是 Imd 信号通路最上游的受体之一。为了研究斯氏按蚊 *Anopheles stephensi* 肽聚糖识别蛋白 PGRP-LC1, 采用 RT-PCR 并结合 RACE 技术克隆斯氏按蚊 *PGRP-LC1* 基因, 通过序列比较分析, 得到两条 cDNA 序列, 其开放阅读框分别为 1 365 bp 和 1 290 bp, 3′非编码区为 320 bp, 5′非编码区为240 bp。将两条 cDNA 分别命名为 *AsPGRP-LC1a*(GenBank 注册号 GU214232)和 *AsPGRP-LC1b*(GenBank 注册号 GU214233)。*AsPGRP-LC1a* 编码 454 个氨基酸, 分子量约为 49.07 kDa; *AsPGRP-LC1b* 编码 429 个氨基酸, 分子量约为 46.3 kDa。*AsPGRP-LC1b* 比 *AsPGRP-LC1a* 少一个长度为 75 bp 的外显子, 该外显子在冈比亚按蚊 *Anopheles gambiae PGRP-LC1* 基因的某些可变剪切形式中也有发现。分别将两个斯氏按蚊 *PGRP-LC1* 基因在冈比亚按蚊细胞系 L3-5 和斯氏按蚊细胞系 MSQ43 中过量表达, 通过双荧光素酶检测系统检测抗菌肽的表达情况, 结果显示克隆得到的 *PGRP-LC1* 基因在两种细胞系中均能够启动 Imd 信号通路, 为进一步研究斯氏按蚊的 Imd 信号通路提供了依据。

关键词: 按蚊; 斯氏按蚊; 疟疾; 肽聚糖识别蛋白; Imd 信号通路; 克隆; 过量表达; 抗菌肽

中图分类号: Q966 **文献标识码:** A **文章编号:** 0454-6296(2010)02-0131-08

Molecular cloning and functional characterization of *PGRP-LC1* in *Anopheles stephensi* (Diptera: Culicidae)

CHEN Yang[1,2], LING Er-Jun[1,*] (1. Research Center for Insect Science, Institute of Plant Physiology and Ecology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200032, China; 2. Graduate University of Chinese Academy of Sciences, Beijing 100080, China)

Abstract: Insect defense against pathogen invasion mainly rely on innate immunity system. Recent research showed that Imd pathway is involved in limiting the number of *Plasmodium berghei* oocysts developing in mosquito midgut, and PGRP-LC1 is one of the receptors of Imd pathway. In order to investigate PGRP-LC1 in *Anopheles stephensi*, we cloned the *PGRP-LC1* gene through a combination of reverse transcription polymerase chain reaction (RT-PCR) and rapid amplification of cDNA ends (RACE) approaches. Two cDNA fragments, with the open reading frame (ORF) of 1 365 bp and 1 290 bp respectively, were obtained. The 3′ untranslated region (UTR) is 320 bp, and the 5′ UTR 240 bp. We named the two sequences as *AsPGRP-LC1a* (GenBank accession no. GU214232) and *AsPGRP-LC1b* (GenBank accession no. GU214233), respectively. *AsPGRP-LC1a* encodes a polypeptide of 454 amino acids with a predicted molecular weight of 49.07 kDa. *AsPGRP-LC1b* encodes a polypeptide of 429 amino acids with a predicted molecular weight of 46.3 kDa. *AsPGRP-LC1b* contains one exon less than *AsPGRP-LC1a*. The exon is 75 bp and present in some of the alternative splicing isoforms of *Anopheles gambiae PGRP-LC1* too. The two PGRP-LC1 genes were over-expressed in both *A. gambiae* cell line L3-5 and *A. stephensi* cell line MSQ43. The expression profile of antimicrobial peptides was monitored by dual luciferase assay system, and the results showed that the *PGRP-LC1* we cloned could activate Imd pathway in both cell lines, which provided basis for further investigations of Imd pathway in *A. stephensi*.

基金项目: 国家"973"计划项目(2007CB513107); 国家自然科学基金面上项目(30970408); 中国高技术研究发展计划("863"计划)项目(2006AA10A119)

作者简介: 陈杨, 男, 1983 年生, 四川成都人, 博士研究生, 研究方向为昆虫天生免疫, E-mail: chenyang83@gmail.com

* 通讯作者 Corresponding author, Tel.: 021-54924155; E-mail: erjunling@sippe.ac.cn

收稿日期 Received: 2009-11-20; 接受日期 Accepted: 2010-01-23

Key words: Anopheline mosquito; *Anopheles stephensi*; malaria; peptidoglycan recognition protein; Imd pathway; cloning; over-expression; antimicrobial peptide

昆虫占地球已知动物物种的90%以上，它们能如此成功地在地球上生存和繁衍，其中很重要一部分原因要归功于其强大的天生免疫系统。和脊椎动物不同，昆虫没有获得性免疫，为了抵御外界微生物和寄生虫的入侵，它们只能依靠天生免疫来完成。肽聚糖识别蛋白(peptidoglycan recognition protein, PGRP)家族就是昆虫天生免疫系统中一个重要的组成部分(Steiner, 2004)。第一个肽聚糖识别蛋白是在家蚕 *Bombyx mori* 中发现的，该蛋白能够结合肽聚糖，从血清中清除该蛋白后，多酚氧化酶级联反应将被抑制(Yoshida *et al*., 1996)。目前在黑腹果蝇 *Drosophila melanogaster* 基因组中共发现了13个肽聚糖识别蛋白基因(Adams *et al*., 2000; Myers *et al*., 2000)。根据大小不同，这些肽聚糖识别蛋白被分为两大类：较短的PGRP-S，主要是分泌型蛋白，较长的PGRP-L主要是跨膜蛋白或细胞内蛋白(Yoshida *et al*., 1996; Werner *et al*., 2000)。

PGRP-LC是PGRP-L家族中的一员。在果蝇中，PGRP-LC负责感受革兰氏阴性细菌的刺激，并激活Imd信号通路(Michel *et al*., 2001; Gottar *et al*., 2002; Choe *et al*., 2005)。研究还发现PGRP-LE能够与PGRP-LC协同作用，共同激活Imd信号通路(Takehana *et al*., 2004)。Imd信号通路被激活以后，将导致NF-κB转录因子Relish入核，启动一系列抗菌肽的表达，如diptericin和cecropin(Choe *et al*., 2005)。Imd信号通路的激活还会导致PGRP-LB的上调，该蛋白能够降解革兰氏阴性细菌的肽聚糖，从而下调Imd信号通路，起到负调控的作用(Zaidman-Remy *et al*., 2006)。

在冈比亚按蚊 *Anopheles gambiae* 中，目前发现了7个肽聚糖识别蛋白基因，其中包括3个 *PGRP-S* 和4个 *PGRP-L*(Christophides *et al*., 2002)。与果蝇中情况相似，冈比亚按蚊 *PGRP-LC* 基因通过可变剪切产生至少3种不同形式，*PGRP-LC1*、*PGRP-LC2* 和 *PGRP-LC3*(Christophides *et al*., 2002)，而这3种形式的可变剪切又可进一步被细分成更多的种类(Lin *et al*., 2007)。在细胞系中运用RNA干扰(RNAi)技术，分别下调 *PGRP-LC* 各种可变剪切形式的基因表达，证明PGRP-LC1对于Imd信号通路的激活至关重要，而过表达PGRP-LC1则会导致抗菌肽基因，如 *Cec1* 和 *Gam1* 等的表达量的大幅度提高(Lin *et al*., 2007)。另外，Imd信号通路也与控制蚊虫中肠中疟原虫卵囊数目相关(Meister *et al*., 2005)。当Imd信号通路被下调时，感染蚊虫的中肠疟原虫卵囊数目会明显增加。因此，作为Imd信号通路最上游的受体之一，PGRP-LC1的功能就显得格外重要。

为了研究斯氏按蚊肽聚糖识别蛋白PGRP-LC1的功能，我们采用RT-PCR扩增并结合RACE技术克隆得到其基因的两种可变剪切形式，并分别在冈比亚按蚊细胞系L3-5和斯氏按蚊细胞系MSQ43中过量表达，通过双荧光素酶检测系统检测抗菌肽的表达情况，结果显示，克隆得到的斯氏按蚊 *PGRP-LC1* 基因在两种细胞系中均能够启动Imd信号通路，导致相应抗菌肽基因表达量的升高。

1 材料和方法

1.1 材料

1.1.1 细胞系：斯氏按蚊细胞系MSQ43(Luna *et al*., 2006)和冈比亚按蚊细胞系L3-5(Zheng *et al*., 1997)为本实验室保存，均使用含10%胎牛血清(FBS)的施耐德培养基(Schneider's *Drosophila* medium, Invitrogen)进行培养，培养温度27℃。

1.1.2 主要试剂：施耐德培养基、胎牛血清和RACE试剂盒、表达质粒pAC5.1/V5-HisB均购自Invitrogen公司；RNA抽提试剂盒、质粒抽提试剂盒和细胞转染试剂盒均购自Qiagen公司；cDNA合成试剂盒购自Toyobo公司；限制性内切酶、PCR Taq酶、DNA marker和T载体购自TaKaRa公司；双荧光素酶检测试剂盒购自Promega公司；用于Western blotting的一抗二抗分别购自Abcam和Chemicon公司；蛋白Marker购自全氏金公司；碱性磷酸酶检测试剂盒购自Bio-Rad公司；引物合成与DNA测序均由Invitrogen公司完成。

1.2 RNA提取和cDNA合成

取1 mL，约 10^7 个MSQ43细胞，用于抽提RNA。总RNA由Qiagen RNeasy试剂盒抽取，并用RNase-free DNAse处理，以去除基因组DNA的污染。取2 μg总RNA按照Toyobo反转录试剂盒说明书反转录合成cDNA。

1.3 斯氏按蚊 *PGRP-LC1* 基因的克隆

通过对已知的冈比亚按蚊 *PGRP-LC1* 和黑腹果蝇 *PGRP-LC* 基因序列的比较分析，设计简并引物MSQF：

5′-AACTCSWCRGAYRTSACGTTY-3′和 MSQR: 5′-CRTCNCCNCCNAYNARAAANT-3′, 扩增斯氏按蚊 *PGRP-LC1* 基因片段。PCR 产物连接 pMD18-T 载体(TaKaRa), 并送测序。测序结果通过 BLAST 分析, 确定是 *PGRP-LC1* 基因片段。

根据得到的基因片段, 设计特异的 3′RACE 引物: 5′-GATATCGCTCATGATCGTTCTGT-3′ 和 5′ RACE 引物 5′-CATGACGATAAAGCTGAGCAA-3′, 配合 Invitrogen RACE 试剂盒扩增斯氏按蚊 *PGRP-LC1* 基因 3′和 5′端序列。再根据获得的 3′和 5′端序列, 设计引物 AsLCNotIF:5′-AATTTGCGGCCGCATGGCACTGGCAAGGAACG-3′和 AsLCXbaIR:5′-GCTCTAGACTAAAATGTTCACTGAAATGCGGC-3′ 扩增 *AsPGRP-LC1* 全长基因。PCR 产物连接 pMD18-T 载体, 并送测序。用 *Not* Ⅰ和 *Xba* Ⅰ从 T 载体上双酶切下基因片段, 并亚克隆到 pAC5. 1/V5-HisB 载体上, 这样, 我们就将斯氏按蚊 *PGRP-LC1* 基因阅读框放到了黑腹果蝇 *Actin 5C* 启动子之后, 并在 3′端连接了 V5 和 His 表达标签。

1.4　目的基因的生物信息学分析

序列同源性比对和相似性搜索用 BLAST 软件进行(http://www. ncbi. nlm. nih. gov/ BLAST/)。蛋白质结构域预测采用 SMART 软件(http://smart. embl-heidelberg. de/)。采用 CLUSTALX 和 MEGA 4. 1软件构建进化树。

1.5　细胞转染

转染所用质粒: 过量表达质粒 pAC5. 1/V5-HisB-AsPGRP-LC1a 和 pAC5. 1/V5-HisB-AsPGRP-LC1b 为本实验构建, pAC5. 1/V5-HisB-AgPGRP-LC1 为本实验室之前构建(Chen *et al*., 2009)。双荧光素酶检测系统所用质粒均为本实验室之前构建(Zheng and Zheng, 2002; Hoa and Zheng, 2007)。pGL3-Cec1, pGL3-Gam1 和 pGL3-Def1 分别为将 *Cec1*, *Gam1* 和 *Def1* 基因的启动子置于萤火虫 *Photinus pyralis* 荧光素酶基因 *luc2* 上游。作为内参的质粒 pAct5C-Renilla 为将黑腹果蝇 *Acting 5C* 启动子置于海参 *Renilla reniformis* 荧光素酶基因 *hRluc* 上游。

转染之前在 24 孔板内铺细胞 L3-5 或 MSQ43, 大约 5×10^5 个细胞/孔, 待细胞密度达到70% ~90%时进行转染。转染使用 Qiagen 转染试剂盒 Effectene。每孔转染质粒包括: 200 ng 过量表达质粒 pAC5. 1/V5-HisB-AsPGRP-LC1a, pAC5. 1/V5-HisB-AsPGRP-LC1b 或 pAC5. 1/V5-HisB-AgPGRP-LC1, 200 ng pGL3-Cec1, pGL3-Gam1 或 pGL3-Def1, 3 ng pAct5C-Renilla。转染 24 h 后, 收集细胞, 用于 Western blotting 和双荧光素酶检测。

1.6　双荧光素酶检测

细胞转染 24 h 后, 使用 Promega 公司 Dual Luciferase Assay 试剂盒检测荧光, 荧光强度由荧光光度计测定。每组实验均重复 3 次, 以保证结果真实可信。每次实验包含 3 次独立的检测。以 *Actin 5C* 启动子启动表达的 *Renilla* 荧光素酶活性为内参, 将各组萤火虫荧光素酶活性代表的抗菌肽启动子活性均一化。

1.7　Western blot

细胞转染 24 h 后, 离心细胞, 取沉淀, 加入 SDS-PAGE 上样缓冲液, 剧烈震荡混匀后, 沸水浴 5 min。样品在 10% 的 SDS-PAGE 中电泳之后, 转膜, 分别以兔源的抗表达标签 V5 的抗体(Abcam, Cambridge, UK)为一抗, 交联碱性磷酸酶的羊抗兔抗体(AP, Chemicon, Temecula, CA)为二抗进行 Western blotting。碱性磷酸酶活性由 Bio-Rad 公司的碱性磷酸酶检测试剂盒检测。

2　结果与分析

2.1　斯氏按蚊 *PGRP-LC1* 基因的克隆和序列分析

RACE 分别获得 *AsPGRP-LC1* 基因 5′和 3′端序列之后, 根据基因序列和表达载体需要设计了正向和反向引物, 用 cDNA 扩增, 获得约 1. 3 kb 条带, 割胶回收, 连接 T 载体, 转化大肠杆菌 *Escherichia coli* TOP10, 挑取 5 个阳性克隆送测序。测序结果显示, 胶回收的 PCR 条带并不单一, 其中包括两条序列, 两条序列之间的差异仅为 75 bp。为了区别, 将较长的序列命名为 *AsPGRP-LC1a*, 较短的序列命名为 *AsPGRP-LC1b*。两条序列的开放阅读框长度分别为 1 365 bp 和 1 290 bp, 蛋白预测分子量分别为 49. 07 kDa 和 46. 3 kDa。cDNA 序列 3′非编码区长度为 320 bp, 5′非编码区为 240 bp。*AsPGRP-LC1a* 和 *AsPGRP-LC1b* 序列均在 GenBank 提交, 注册号分别为 GU214232 和 GU214233。

通过 BLAST 分析发现, *AsPGRP-LC1a* 比 *AsPGRP-LC1b* 多出一个 75 bp 的外显子, 该外显子在与冈比亚按蚊 *PGRP-LC1* 基因的某些可变剪切形式中也存在, 说明 *AsPGRP-LC1a* 和 *PGRP-LC1b* 是同一个基因在转录时可变剪切产生的不同形式。*AsPGRP-LC1a* 和冈比亚按蚊 *PGRP-LC1* 基因比对显示, 两个基因在核苷酸序列和蛋白序列水平上分别存在 82% 和 72% 的相似性。*AsPGRP-LC1a* 的 cDNA 全长序列见图 1。

```
     TGCTTTCTTCGGTGGCGTTTCTTCGATAAAACTTCAAACCAGTGACTCACCCAGCCAGTT
     TATTTTTACAGACTTACTTCGCCCGTTTTCAACGTTCAGCGGGGGAAGGGAAGGAACATT
     CTTCCGCCAAGATGCCTCTCGATGTTTGCTGGAGATAATAGTATCAACAACTTAAGCAGA
     TAGCTTTCGGAATATCTGACTAGCTGACCACCGACAACCCGCACCATTACGTACCAAACG
   1 M  A  L  A  R  N  A  Q  S  I  E  N  E  A  P  F  K  T  G  G
   1 ATGGCACTGGCAAGGAACGCCCAATCGATAGAAAATGAAGCACCATTCAAGACCGGAGGA
  21 Y  S  T  T  Q  P  P  P  K  G  G  I  D  T  V  A  G  G  V  N
  61 TACAGCACGACACAGCCGCCACCGAAAGGTGGGATCGATACGGTTGCCGGAGGAGTAAAC
  41 S  P  S  T  V  T  S  T  T  T  D  S  T  R  G  S  L  S  N  I
 121 AGCCCGAGTACCGTGACCAGCACCACAACCGACAGCACCAGAGGCAGTCTCTCCAATATC
  61 S  T  S  S  I  K  R  T  D  N  A  C  D  D  D  S  S  A  F  D
 181 TCGACGTCATCCATCAAACGGACAGACAACGCCTGCGATGACGATTCGAGTGCCTTCGAT
  81 S  S  S  S  N  G  G  S  D  T  E  C  E  D  D  T  I  K  R  A
 241 AGCAGCAGCAGCAACGGTGGCAGCGATACCGAGTGTGAGGATGATACCATCAAGCGCGCC
 101 I  D  R  I  P  G  A  L  V  P  G  E  A  R  V  L  P  N  G  N
 301 ATTGACCGTATCCCGGGTGCGCTGGTACCCGGGGAAGCACGCGTACTCCCGAACGGGAAT
 121 L  K  V  G  A  E  A  I  S  P  A  V  R  P  N  P  G  Q  P  S
 361 CTTAAGGTTGGCGCTGAAGCGATCAGTCCCGCAGTGCGGCCCAACCCGGGCCAACCCTCC
 141 A  I  G  A  I  A  V  H  N  S  S  D  I  T  F  G  N  K  T  Y
 421 GCCATCGGGGCGATCGCGGTGCACAACTCGTCCGACATCACGTTCGGCAACAAAACGTAC
 161 I  K  G  Q  V  V  I  K  N  I  Y  Q  D  R  K  Q  G  A  A  N
 481 ATTAAGGGGCAGGTGGTGATAAAGAACATCTACCAGGATCGTAAGCAGGGGGCCGCCAAC
 181 Q  G  Y  S  V  H  E  N  D  T  T  D  T  S  A  T  K  K  V  P
 541 CAGGGATACAGTGTGCATGAAAACGATACTACCGACACCAGTGCTACTAAAAAAGTACCT
 201 P  S  A  S  I  P  P  E  P  R  T  W  Q  S  S  L  K  T  I  I
 601 CCTTCAGCGTCGATACCTCCGGAGCCACGCACCTGGCAATCTAGTCTAAAGACAATCATC
 221 K  D  K  P  L  L  S  F  I  V  M  I  S  L  M  I  V  L  C  A
 661 AAAGACAAACCGTTGCTCAGCTTTATCGTCATGATATCGCTCATGATCGTTCTGTGTGCC
 241 I  V  A  V  I  S  I  L  T  A  S  G  K  A  R  I  R  P  P  L
 721 ATCGTAGCGGTCATATCCATCCTCACCGCGTCAGGTAAAGCTAGAATAAGGCCACCGCTT
 261 G  D  G  D  D  N  R  P  N  I  P  Q  D  K  D  I  E  D  D  L
 781 GGCGACGGCGACGACAATCGTCCCAATATCCCACAGGACAAGGACATAGAGGACGACCTG
 281 F  P  D  P  R  P  L  R  L  V  T  R  T  E  W  L  A  Q  P  P
 841 TTCCCTGATCCTAGGCCATTGCGGCTGGTGACGAGAACCGAGTGGCTTGCTCAACCGCCT
 301 R  E  E  L  T  D  L  K  L  P  V  H  N  V  I  I  A  H  T  A
 901 CGCGAAGAACTGACTGATCTTAAGCTACCGGTGCACAACGTCATCATTGCTCACACTGCC
 321 T  E  G  C  T  T  Q  Q  A  C  K  A  L  V  Q  S  I  Q  R  Y
 961 ACCGAGGGGTGCACTACTCAGCAAGCATGCAAAGCACTAGTGCAATCCATTCAAAGGTAC
 341 H  V  N  K  T  K  Y  G  D  I  G  Y  N  F  L  I  G  G  D  A
1021 CACGTAAACAAAACCAAGTACGGTGATATCGGGTACAATTTTTTGATCGGTGGCGATGCA
 361 F  I  Y  E  G  R  G  W  L  K  M  G  A  H  T  K  D  Y  N  V
1081 TTTATCTACGAGGGCCGGGGATGGCTCAAGATGGGTGCCCACACCAAGGACTACAACGTC
 381 K  S  H  G  I  A  F  I  G  N  Y  E  T  V  D  R  P  S  E  E
1141 AAAAGCCACGGGATTGCCTTCATCGGTAACTACGAAACTGTCGATCGACCGTCCGAAGAA
 401 Q  I  E  Q  L  E  L  L  L  R  N  G  T  D  G  G  W  L  G  S
1201 CAAATCGAACAGCTGGAGCTGTTGCTTCGCAACGGTACCGATGGTGGGTGGCTTGGGTCG
 421 D  Y  R  L  F  G  A  S  Q  L  K  S  T  I  S  P  G  K  L  L
1261 GACTACCGGTTGTTCGGGGCAAGTCAGCTTAAATCCACCATCAGTCCGGGGAAATTGCTG
 441 M  E  K  L  R  L  L  P  H  F  S  E  H  F  *
1321 ATGGAAAAGTTACGTCTGCTGCCGCATTTCAGTGAACATTTTTAAAGAGGCGAAAAAGTG
     ATCGAGAGTGCACAGGATCGCGTAAGACTGAAGGGGGGGCCTCGTGCTGTAAAACGGTCG
     ATGCAACCGTACGATAAATATATTTATGGTGGGAAACATCGTTTTAAAAATACTTTTAAG
     AATAAGGATCCTTTCAATGCTTCGCAAATGTAGTGAGTGTGATCGTGGACGAACGAACTG
     CTTCCGCAATATTCTTAACATAGTTGGATATCCGTATCCGTATGATACGTTACGTTTATG
     ACATCAATAAAATACTTCTTTTAAACACATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
     AAAAA
```

图 1 斯氏按蚊 *AsPGRP-LC1a* 基因的编码区核苷酸序列及推导的氨基酸序列(包括 5′和 3′非编码区)

Fig. 1 Nucleotide and the deduced amino acid sequences of encoding region of *AsPGRP-LC1a* in *Anopheles stephensi* (including 5′ and 3′ UTRs)

黑色部分是推测的 *AsPGRP-LC1a* 跨膜区，灰色部分是 *AsPGRP-LC1a* 比 *AsPGRP-LC1b* 多出的 75 bp 序列，方框内是起始密码子和终止密码子。
The putative transmembrane domain sequences of *AsPGRP-LC1a* are shown in black parts. The 75 bp sequences that *AsPGRP-LC1a* differs from *AsPGRP-LC1b* are shown in grey parts. Start codon and stop codon are boxed.

运用 CLUSTAL X 和 MEGA4.1 软件，采用 Neighbor-joining 法构建进化树(图 2)。结果显示各蚊种包括斯氏按蚊、冈比亚按蚊、埃及伊蚊 *Aedes aegypti* 和致倦库蚊 *Culex quiquefasciatus* 的 PGRP-LC 聚在一起，其中斯氏按蚊和冈比亚按蚊进化关系较近，埃及伊蚊和致倦库蚊进化关系较近。两个 AsPGRP-LC1 与 AgPGRP-LC1 聚到一起，进一步证明我们获得的序列确实是 PGRP-LC1。另外烟草天蛾 *Manduca sexta* 和家蚕的 PGRP 聚在一起，黑腹果蝇和意大利蜜蜂 *Apis mellifera* 的 PGRP 聚在一起。

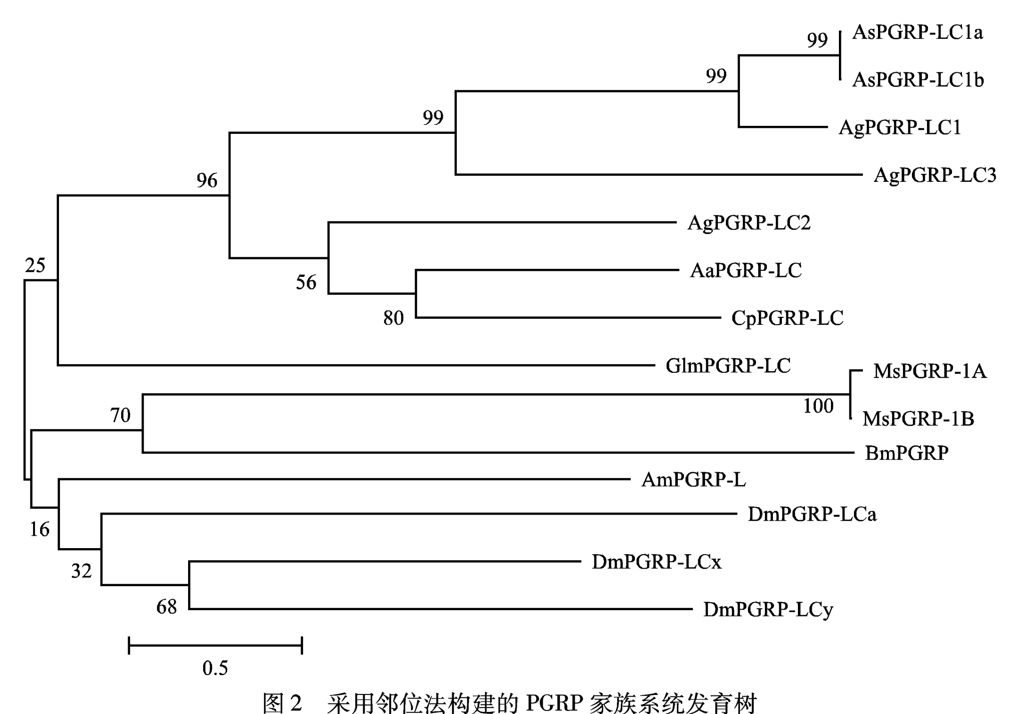


图 2　采用邻位法构建的 PGRP 家族系统发育树

Fig. 2　A phylogenetic tree of PGRP family members constructed with the neighbour-joining method

生物种名及各序列的 GenBank 注册号如下 Species names and the GenBank accession numbers are as follows：斯氏按蚊 *Anopheles stephensi*：AsPGRP-LC1a（GU214232），AsPGRP-LC1b（GU214233）；冈比亚按蚊 *Anopheles gambiae*：AgPGRP-LC1（XM_314103），AgPGRP-LC2（XM_558599），AgPGRP-LC3（XM_558600）；埃及伊蚊 *Aedes aegypti*：AaPGRP-LC（XM_001649201.1）；致倦库蚊 *Culex quiquefasciatus*：CpPGRP-LC（XM_001848006.1）；烟草天蛾 *Manduca sexta*：MsPGRP-1A（AF413068），MsPGRP-1B（AF413061）；家蚕 *Bombyx mori*：BmPGRP（AF441723）；黑腹果蝇 *Drosophila melanogaster*：DmPGRP-LCa（NM_140041）；DmPGRP-LCx（NM_168324），DmPGRP-LCy（NM_206308）；舌蝇 *Glossina morsitans*：GlmPGRP-LC（DQ307161）；意大利蜜蜂 *Apis mellifera*：AmPGRP-L（XM_392452）.

2.2　*AsPGRP-LC1a* 和 *AsPGRP-LC1b* 在细胞系中表达

将 *AsPGRP-LC1a* 和 *AsPGRP-LC1b* 分别连入表达载体 pAC5.1/V5-HisB，然后转染斯氏按蚊细胞系 MSQ43。24 h 后取全细胞进行 Western blotting，结果如图 3。从图中可以看到，在 40～60 kDa 范围内有明显条带，与理论大小相符。证明我们克隆得到的 *AsPGRP-LC1a* 和 *AsPGRP-LC1b* 均能在细胞系中正常表达。由于两个蛋白之间分子量相差很小，所以在 SDS-PAGE 中也无法区分。

2.3　*AsPGRP-LC1a* 和 *AsPGRP-LC1b* 在细胞系中过量表达对 *Cec1* 表达量的影响

我们分别将冈比亚按蚊 *AgPGRP-LC1*（Chen *et al.*，2009）及斯氏按蚊 *AsPGRP-LC1a* 和 *AsPGRP-LC1b* 在 L3-5 和 MSQ43 两种细胞系中表达，再用双荧光素酶检测法检测 Imd 信号通路下游抗菌肽基因 *Cec1* 表达情况。结果显示在冈比亚按蚊细胞系 L3-5（图 4：A）和斯氏按蚊细胞系 MSQ43（图 4：B）中，过量表达 *AsPGRP-LC1a* 和 *AsPGRP-LC1b* 均能够导致 *Cec1* 表达量升高，且 *Cec1* 表达量升高幅度相当。另外，在两种细胞系中，过量表达 *AsPGRP-LC1a* 和 *AsPGRP-LC1b* 后 *Cec1* 表达量均稍高于过量表达冈比亚按蚊 *AgPGRP-LC1* 后的情况，显示 AsPGRP-LC1a 和 AsPGRP-LC1b 对于 Imd 信号通路的激活能力稍强于 AgPGRP-LC1，其中原因并不清楚。

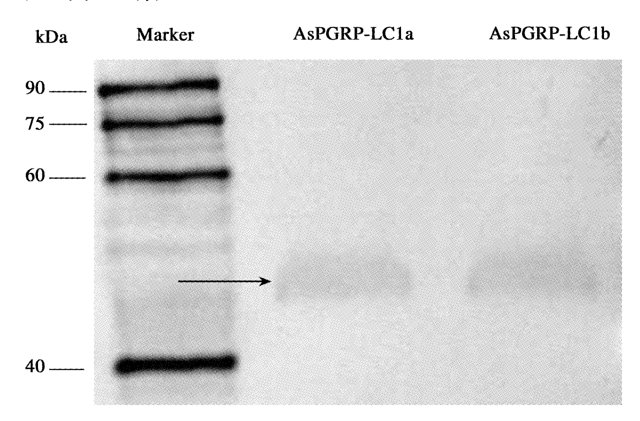


图 3　斯氏按蚊 AsPGRP-LC1a 和 AsPGRP-LC1b 在细胞内过量表达后的 Western blotting 检测结果

Fig. 3　Analysis of AsPGRP-LC1 and AsPGRP-LC1b expressed in the cell by Western blotting

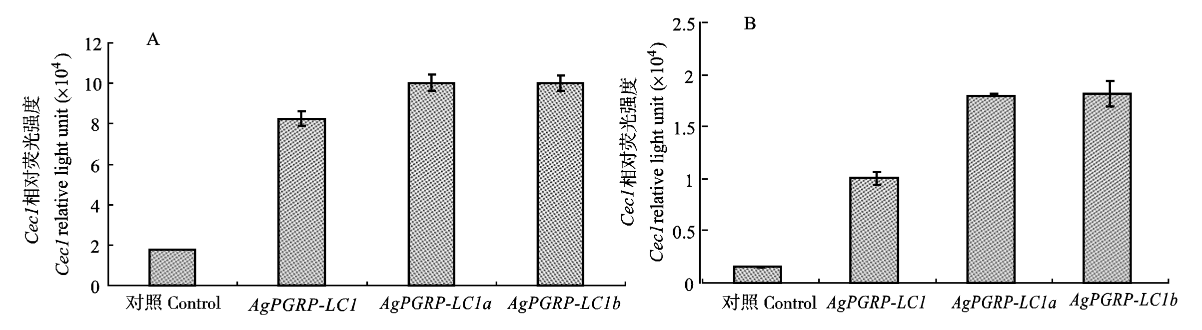


图 4 冈比亚按蚊 *AgPGRP-LC1*、斯氏按蚊 *AsPGRP-LC1a* 和 *AsPGRP-LC1b* 分别在细胞系 L3-5 和 MSQ43 中过量表达后其细胞中 *Cec1* 基因的相对表达强度

Fig. 4 Relative expression level of *Cec1* in cell lines L3-5 and MSQ43 after over-expressing *AgPGRP-LC1*, *AsPGRP-LC1a* and *AsPGRP-LC1b* respectively

A 和 B 分别是在细胞系 L3-5 和 MSQ43 中过量表达的情况;误差棒表示至少 3 次独立实验的标准偏差。A and B are the situations in cell lines L3-5 and MSQ43, respectively. Error bars indicate standard deviations from at least three independent assays.

2.4 *AsPGRP-LC1a* 和 *AsPGRP-LC1b* 在细胞系中过量表达对 *Gam1* 和 *Def1* 表达量的影响

为了进一步验证 *AsPGRP-LC1a* 和 *AsPGRP-LC1b* 的功能，我们又检测了另外两个 Imd 信号通路调控的抗菌肽基因 *Gam1* 和 *Def1* 的表达量，得到与 *Cec1* 类似的结果(图 5)。*Def1* 与另外两种抗菌肽相比，表达量相对较高，不管在有没有过量表达 *PGRP-LC1* 的情况下，其表达量均高于 *Gam1* 和 *Cec1*。而 *Gam1* 和 *Cec1* 的表达量相当。

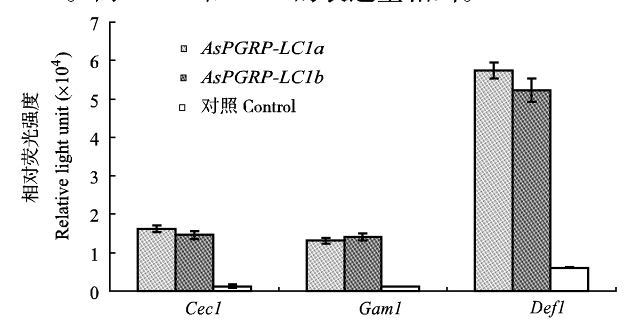


图 5 斯氏按蚊 *AsPGRP-LC1a* 和 *AsPGRP-LC1b* 在细胞系 MSQ43 中过量表达后 *Cec1*, *Gam1* 和 *Def1* 的表达情况

Fig. 5 Relative expression level of *Cec1*, *Gam1 and Def1* in cell line MSQ43 after over-expressing *AsPGRP-LC1a* and *AsPGRP-LC1b*, respectively

误差棒表示至少 3 次独立实验的标准偏差。Error bars indicate standard deviations from at least three independent assays.

3 讨论

本文克隆得到两条斯氏按蚊 *PGRP-LC1* 基因，被分别命名为 *AsPGRP-LC1a* 和 *AsPGRP-LC1b*。*AsPGRP-LC1a* 比 *AsPGRP-LC1b* 多出一个长度为 75 bp的外显子，该外显子在冈比亚按蚊 *PGRP-LC1* 中也有发现。通过与冈比亚按蚊 *PGRP-LC1* 序列比对，发现斯氏按蚊 *PGRP-LC1* 与其在核苷酸水平和蛋白水平相似性分别达到 82% 和 72%，显示出这两个物种之间很近的亲缘关系。

之前的研究已经证明，在冈比亚按蚊细胞系 4A3a(Lin *et al*., 2007)和 L3-5(Chen *et al*., 2009)中过量表达 PGRP-LC1 能导致 Imd 信号通路下游的抗菌肽基因表达量上调。将克隆得到的两个基因分别在冈比亚按蚊和斯氏按蚊细胞系中过量表达，结果显示克隆得到的两条序列均能很好地启动 Imd 信号通路，导致该通路调控的抗菌肽基因表达量提高。而且，对于 Imd 信号通路的激活程度，两条基因之间并没有太大差别。因此，*AsPGRP-LC1a* 多出的 75 bp 的外显子的功能并不清楚。该外显子位于细胞膜外结构域，因此我们猜想，这种可变剪切的目的在于扩大受体识别外源物的范围。在 MSQ43 细胞系中，*Def1* 的表达量不管在有没有过量表达 PGRP-LC1 的情况下，均高于 *Gam1* 和 *Cec1*。这应该与抗菌肽自身启动子特性有关。之前的研究发现，去除细胞外结构域的 PGRP-LC 依然能够启动 Imd 信号通路(Choe *et al*., 2005)。而且，剪切掉细胞外结构域的 PGRP-LC 对于 Imd 信号通路的启动能力反而更强(Schmidt *et al*., 2008)，且细胞膜外结构域越短，其活性越强(Chen *et al*., 2009)。因此，Schmidt 等(2008)提出假说，Imd 信号通路也存在和 Toll 信号通路类似的激活方式，即通过剪切细胞膜上的受体来实现。要证明这个假说，还有很多很多问题需要解决，如是由什么蛋白来对 PGRP-LC 进行剪切，这个或这些蛋白又是在什么情况下会具有剪切的活

性等等。

另外，在对果蝇的研究中发现，对于 Imd 信号通路的激活，除了 PGRP-LC 在起作用之外，还有其他 PGRP 蛋白在起作用，如 PGRP-LE 和 PGRP-LF。PGRP-LE 与 PGRP-LC 协同作用，启动 Imd 信号通路（Kaneko *et al*., 2006），而 PGRP-LF 则起到负调控的作用（Persson *et al*., 2007; Maillet *et al*., 2008）。但是这些蛋白在冈比亚按蚊的基因组中都没有找到对应的同源蛋白。是不是按蚊中就不存在更多的调控蛋白呢？还是存在另外的调控机制？这些都是亟待解决的问题。之前对于冈比亚按蚊的研究也暗示，在 Imd 信号通路的激活过程中可能存在某些类似细胞因子的蛋白在起作用（Lin *et al*., 2007）。但是到目前为止，也没有任何的发现。

按蚊是疟疾、丝虫病等传染性疾病的重要媒介，遗传改造蚊媒，加强蚊媒自身的免疫能力，抵御寄生虫的入侵，从而切断寄生虫的传播通道，是 WHO 早在上世纪 90 年代初就提出的思路。但是，要实现这一目标，首先就需要对蚊媒的天生免疫系统有全面的了解。虽然近年来在昆虫天生免疫方面有很多新的发现，但是依然有很多问题需要解答。目前的研究已经显示 Imd 信号通路与疟原虫的感染强度存在密切的关系，因此，对 Imd 信号通路在细胞膜上的受体 PGRP-LC 的功能以及激活方式进行深入研究就显得更为重要。

致谢　感谢中国科学院上海生命科学研究院植物生理生态研究所郑良彪教授在研究过程中给予的帮助和指导。

参 考 文 献（References）

Adams MD, Celniker SE, Holt RA, *et al*., 2000. The genome sequence of *Drosophila melanogaster*. *Science*, 287(5 461): 2 185 – 2 195.

Chen Y, Ling EJ, Weng ZH, 2009. Functional characterization of PGRP-LC1 of *Anopheles gambiae* through deletion and RNA interference. *Insect Science*, 16(6): 443 – 453.

Choe KM, Lee H, Anderson KV, 2005. *Drosophila* peptidoglycan recognition protein LC (PGRP-LC) acts as a signal-transducing innate immune receptor. *Proc. Natl. Acad. Sci. USA*, 102(4): 1 122 – 1 126.

Christophides GK, Zdobnov E, Barillas-Mury C, Birney E, Blandin S, Blass C, Brey PT, Collins FH, Danielli A, Dimopoulos G, Hetru C, Hoa NT, Hoffmann JA, Kanzok SM, Letunic I, Levashina EA, Loukeris TG, Lycett G, Meister S, Michel K, Moita LF, Muller HM, Osta MA, Paskewitz SM, Reichhart JM, Rzhetsky A, Troxler L, Vernick KD, Vlachou D, Volz J, von Mering C, Xu J, Zheng L, Bork P, Kafatos FC, 2002. Immunity-related genes and gene families in *Anopheles gambiae*. *Science*, 298(5 591): 159 – 165.

Gottar M, Gobert V, Michel T, Belvin M, Duyk G, Hoffmann JA, Ferrandon D, Royet J, 2002. The *Drosophila* immune response against Gram-negative bacteria is mediated by a peptidoglycan recognition protein. *Nature*, 416(6 881): 640 – 644.

Hoa NT, Zheng LB, 2007. Functional characterization of the NF-kappa B transcription factor gene REL2 from *Anopheles gambiae*. *Insect Science*, 14(3): 175 – 184.

Kaneko T, Yano T, Aggarwal K, Lim JH, Ueda K, Oshima Y, Peach C, Erturk-Hasdemir D, Goldman WE, Oh BH, Kurata S, Silverman N, 2006. PGRP-LC and PGRP-LE have essential yet distinct functions in the *Drosophila* immune response to monomeric DAP-type peptidoglycan. *Nat. Immunol*., 7(7): 715 – 723.

Lin H, Zhang LM, Luna C, Hoa NT, Zheng LB, 2007. A splice variant of PGRP-LC required for expression of antimicrobial peptides in *Anopheles gambiae*. *Insect Science*, 14(3): 185 – 192.

Luna C, Hoa NT, Lin H, Zhang L, Nguyen HL, Kanzok SM, Zheng L, 2006. Expression of immune responsive genes in cell lines from two different *Anopheline* species. *Insect Mol. Biol*., 15(6): 721 – 729.

Maillet F, Bischoff V, Vignal C, Hoffmann J, Royet J, 2008. The *Drosophila* peptidoglycan recognition protein PGRP-LF blocks PGRP-LC and IMD/JNK pathway activation. *Cell Host and Microbe*, 3(5): 293 – 303.

Meister S, Kanzok SM, Zheng XL, Luna C, Li TR, Hoa NT, Clayton JR, White KP, Kafatos FC, Christophides GK, Zheng L, 2005. Immune signaling pathways regulating bacterial and malaria parasite infection of the mosquito *Anopheles gambiae*. *Proc. Natl. Acad. Sci. USA*, 102(32): 11 420 – 11 425.

Michel T, Reichhart JM, Hoffmann JA, Royet J, 2001. *Drosophila* Toll is activated by Gram-positive bacteria through a circulating peptidoglycan recognition protein. *Nature*, 414 (6 865): 756 – 759.

Myers EW, Sutton GG, Delcher AL, Dew IM, Fasulo DP, Flanigan MJ, Kravitz SA, Mobarry CM, Reinert KH, Remington KA, Anson EL, Bolanos RA, Chou HH, Jordan CM, Halpern AL, Lonardi S, Beasley EM, Brandon RC, Chen L, Dunn PJ, Lai Z, Liang Y, Nusskern DR, Zhan M, Zhang Q, Zheng X, Rubin GM, Adams MD, Venter JC, 2000. A whole-genome assembly of *Drosophila*. *Science*, 287(5 461): 2 196 – 2 204.

Persson C, Oldenvi S, Steiner H, 2007. Peptidoglycan recognition protein LF: a negative regulator of *Drosophila* immunity. *Insect Biochem. Mol. Biol*., 37(12): 1 309 – 1 316.

Schmidt RL, Trejo TR, Plummer TB, Platt JL, Tang AH, 2008. Infection-induced proteolysis of PGRP-LC controls the IMD activation and melanization cascades in *Drosophila*. *FASEB J*., 22 (3): 918 – 929.

Steiner H, 2004. Peptidoglycan recognition proteins: on and off switches for innate immunity. *Immunol. Rev*., 198: 83 – 96.

Takehana A, Yano T, Mita S, Kotani A, Oshima Y, Kurata S, 2004. Peptidoglycan recognition protein (PGRP)-LE and PGRP-LC act synergistically in *Drosophila* immunity. *EMBO J*., 23(23): 4 690 – 4 700.

Werner T, Liu G, Kang D, Ekengren S, Steiner H, Hultmark D, 2000. A family of peptidoglycan recognition proteins in the fruit fly *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA*, 97(25): 13 772 - 13 777.

Yoshida H, Kinoshita K, Ashida M, 1996. Purification of a peptidoglycan recognition protein from hemolymph of the silkworm, *Bombyx mori*. *J. Biol. Chem.*, 271(23): 13 854 - 13 860.

Zaidman-Remy A, Herve M, Poidevin M, Pili-Floury S, Kim MS, Blanot D, Oh BH, Ueda R, Mengin-Lecreulx D, Lemaitre B, 2006. The *Drosophila* amidase PGRP-LB modulates the immune response to bacterial infection. *Immunity*, 24(4): 463 - 473.

Zheng L, Cornel AJ, Wang R, Erfle H, Voss H, Ansorge W, Kafatos FC, Collins FH, 1997. Quantitative trait loci for refractoriness of *Anopheles gambiae* to *Plasmodium cynomolgi* B. *Science*, 276 (5 311): 425 - 428.

Zheng XL, Zheng L, 2002. Genomic organization and regulation of three cecropin genes in *Anopheles gambiae*. *Insect Mol. Biol.*, 11(6): 517 - 525.

（责任编辑：赵利辉）